Sampling and Statistical Considerations for Hydroacoustic Surveys Used in Estimating Abundance of Forage Fishes in Reservoirs

J. CHRISTOPHER TAYLOR*
North Carolina State University, Department of Zoology, Center for Marine Sciences and Technology, 303 College Circle, Morehead City, North Carolina 28557, USA

JESSICA S. THOMPSON
North Carolina State University, Department of Zoology, Box 7617, Raleigh, North Carolina 27695, USA

PETER S. RAND
The Wild Salmon Center, 721 NW Ninth Avenue, Suite 280, Portland, Oregon 97209, USA

MONTSERRAT FUENTES
North Carolina State University, Department of Statistics, Box 8203, Raleigh, North Carolina 27695, USA

Abstract.—An understanding of the spatial distribution of forage fish resources is required to make informed fishery management decisions. We used mobile hydroacoustics to assess the distribution and abundance of forage fish in Badin Lake, a reservoir in central North Carolina. By sampling a series of cross-channel and longitudinal transects and analyzing the data using geostatistics, we characterized both large- and small-scale spatial patterns in forage fish density. Forage fish were observed in higher densities in upstream regions of the reservoir and were seen only in surface waters during July 2000 owing to the existence of a strong thermo–oxycline and in two layers (surface and near bottom) during mixed conditions in December 2001. We observed differences in the scale of patchiness (200–700 m) in forage fish distribution depending on the region of the reservoir where sampling took place, and we infer that these patterns are governed by prevailing limnological conditions. Modeling the spatial variation in the acoustic data using geostatistics resulted in similar average densities (July 2000: \(0.56 \pm 0.28\) [mean \pm SD] fish/m\(^2\); December 2001: \(0.57 \pm 0.49\) fish/m\(^2\)) and improvements in the precision of abundance estimates based on approximated variance (July 2000: \(7.05 \times 10^6 \pm 8.30 \times 10^5\) fish; December 2001: \(7.07 \times 10^6 \pm 6.10 \times 10^5\) fish) when compared with arithmetic averaging and extrapolation (July 2000: \(0.60 \pm 0.61\) fish/m\(^2\), \(7.5 \times 10^6 \pm 4.61 \times 10^5\) fish; December 2001: \(0.60 \pm 0.69\) fish/m\(^2\), \(8.00 \times 10^6 \pm 6.34 \times 10^5\) fish). We found that sampling along longitudinal transects is a more efficient way to characterize the spatial patterns of forage fish distribution and to estimate systemwide abundance and biomass than using data collected with both a longitudinal and cross-channel sampling design in this system.

The management of sport fisheries in inland reservoirs is improved through a knowledge of ecosystem processes, particularly the interactions between piscivorous fishes and their forage base (Ney 1981, 1990; Noble 1981). Reaching this goal requires an informed understanding of how organisms are distributed over the broad range of spatial scales at which these interactions occur. Much of the research effort in reservoir fishery studies has focused on large-scale spatial patterns, with an emphasis on physical and biological gradients (e.g., temperature, water flow, turbidity, conductivity, chlorophyll) typically oriented along the longitudinal axis of the water body (Tisa et al. 1987; Tisa and Ney 1991; Michaletz and Gale 1999). Less attention has been focused on smaller scale spatial patterns in these ecosystems (Schael et al. 1995). These smaller scale dynamics can give rise to spatial structuring in fish communities that undoubtedly have influence on higher order ecosystem pro-
processes, including predator–prey and trophic dynamics (Pitcher 1986; Nero and Magnuson 1989; Rose and Leggett 1990; Schael et al. 1995). Efforts to understand the spatial organization of aquatic organisms over a range of spatial scales is critical to resolving trophic interactions that play a key role in mediating the growth and survival of species of interest to managers (Rose and Leggett 1990; Legendre 1993; Mason and Brandt 1999).

Reservoir fisheries managers have employed a variety of techniques to quantify the abundance of forage fish resources, typically on the scale of the entire system or large regions within the system. Assessments of reservoir prey species such as shad *Dorosoma* spp. have relied on the use of active and passive fishing gears such as gill nets, trawls, electrofishing, and toxicants to produce reservoir-wide abundance and biomass estimates for use in sport fish stocking and management strategies (e.g., Noble 1981; Van Den Avyle et al. 1995). However, without significant man-hour investments, these methods generally lack the spatial resolution that is required to rigorously address how piscivores might perceive and select different habitats within the ecosystem, and the ecological consequences that would result from the degree of overlap in the distribution of forage fishes and piscivores.

For nearly 15 years, fisheries hydroacoustics has been employed on many reservoirs throughout the southeastern USA to assess the pelagic forage fish communities (e.g., Degan and Wilson 1995; Vondracek and Degan 1995; Cyterski et al. 2003). Hydroacoustics provides a repeatable, noninvasive method of collecting high-resolution (submeter scale), continuous data along transects in three dimensions (MacLennan and Simmonds 1992). Hydroacoustic sampling programs in reservoirs are varied in their sampling protocol and data analysis procedures. Typical transect sampling designs include single paths following the main channel of the reservoir, a single transect that zigzags from shore to shore, or several parallel transects that run perpendicular to the axis of the reservoir (Jolly and Hampton 1990). Using any of these transect designs results in data that are typically autocorrelated. Design-based abundance estimates are calculated from these data using a block averaging of depth-integrated (two-dimensional) data without regard for spatial variation or to minimize the statistical complications associated with autocorrelated data (Cliff and Ord 1981). Conclusions from work conducted by Vondracek and Degan (1995) suggest the best approach in temperate reservoirs is a stratified random design of transects that incorporate potential large-scale gradients of fish density within the reservoir. The authors further suggest that block averages of transects that are 150–200 m long minimize complications associated with spatial (serial) correlation and reduce the sampling coefficient of variation (100 · SD/mean; Vondracek and Degan 1995). These design-based approaches to abundance estimates require randomized (or similar) sampling design that can be inefficient with respect to the time spent in transit between transects. An improvement to these methods would take advantage of the continuous nature of hydroacoustic survey data. Geostatistics, and specifically kriging, is a model-based approach that does not require a prescribed randomized sampling plan and implicitly models both large- and small-scale spatial variability (Rivoirard et al. 2000).

The objectives of this paper are to (1) document small- and large-scale spatial variation in forage fish density distribution in a reservoir using hydroacoustic data; (2) assess the performance of arithmetic extrapolation and geostatistical modeling, two analytical methods for estimating the systemwide abundance and biomass of forage fishes in a reservoir; (3) compare the performance of geostatistical models for estimating abundance using different subsets of the data set representing different sampling approaches; (4) utilize these geostatistical models to reconstruct density distributions of forage fishes at scales relevant to the spatial processes that are the current focus of predator–prey studies in reservoir systems; and (5) resolve patterns in the vertical layering of the prey fish community and how these patterns may be masked by depth-integrated approaches to characterizing spatial patterns of forage species in reservoirs.

**Methods**

**Study Site**

Badin Lake is a 2,100-ha reservoir impounded in 1917 and located on the Yadkin River in the central Piedmont region in North Carolina, USA (Figure 1). It is a warm, monomictic water body with an average depth of 21 m and a maximum depth of 54 m. The reservoir is generally stratified from June to early September each year and can become anoxic below 10 m as early as July (J. Thompson, unpublished data). Threadfin shad *D. petenense* and blueback herring *Alosa aestivalis* dominate the forage fish community in the pelagic
environment, while gizzard shad *D. cepedianum* comprise an additional component of the prey fish community (J. Thompson, unpublished data).

**Survey Design**

Hydroacoustic surveys were conducted on 10–11 July 2000 and 5–6 December 2001. All sampling was conducted at night when clupeids disaggregate (Schael et al. 1995; Freon and Misund 1999). For the purposes of this paper, the reservoir was divided into three sampling regions (Figure 1). Two transect patterns were sampled on each occasion: a longitudinal transect that was positioned along the midchannel of each of the three regions, and a series of cross-channel transects positioned within each basin. Temperature and dissolved oxygen profiles were obtained either just prior to or following each acoustic survey to provide a measure of available habitat under conditions where thermal and oxygen stratification can restrict fishes to the epilimnion.

Assigning hydroacoustic data to species can be difficult in systems with high diversity. However, most reservoirs in the southeastern USA are dominated by only a few species in the pelagic zone. Purse-seine sampling concurrent with the hydroacoustic survey resulted in 97% threadfin shad and 3% gizzard shad in July 2000, and a mixture of threadfin shad (48%), gizzard shad (9%), and blueback herring (43%) in December 2001 (all values expressed as a percentage of catch by number; J. Thompson, unpublished data). For the purposes of this paper, we grouped all forage species into a single category for assessments of abundance and biomass. We feel this approach is reasonable since an analysis of stomach content data do not indicate species selectivity by striped bass *Morone saxatilis*, a dominant pelagic piscivore in this system (J. Thompson, unpublished data).

**Data Collection**

We used a 200-kHz, Model 241 split-beam echosounder (Hydroacoustic Technology, Inc. [HTI], Seattle, Washington) fitted with either one or two split-beam transducers. The transducer(s) were mounted on a pole that was fixed to the port side of a small boat. For the sampling on July 2000, we used a single elliptical (4° × 10° nominal beam dimensions) transducer. Ping rate was 5 pings/s, pulse width was 0.18 ms, and vessel speed was approximately 1.8 m/s. Each transect was covered twice, first with the elliptical transducer oriented sideward for collecting data in the top 2 m of the water column and from 1 to 30 m perpendicular to the path of the boat. This was immediately followed by retracing the same transect with the elliptical transducer oriented downward for collecting data in the top 2 m of the water column and from 1 to 30 m perpendicular to the path of the boat. This was immediately followed by retracing the same transect with the elliptical transducer oriented downward for collecting data in the top 2 m of the water column and from 1 to 30 m perpendicular to the path of the boat. For sampling on December 2001, the elliptical transducer was oriented sideward and a second circular transducer (15° nominal beam dimension) was oriented downward. Both transducers were mounted to the same pole affixed to the port side of the boat. Data in the top 2 m and from 2 m to the bottom were collected along a transect by the sideward and downward transducers, respectively. Acoustic pings were sent and echoes received by each transducer through a multiplexing sequence controlled by the echosounder. The ping rate was set at 5 pings/s for each transducer. For both transducer arrangements, returning acoustic signals were simultaneously adjusted for 40-log R and 20-log R time-varied gain for split-beam and echo integration processing, respectively. Each sampling occasion was preceded by an in situ calibration using a tungsten–carbide reference sphere of known target strength placed greater than 10 m from the transducer face. Gain parameters were adjusted accordingly based on calibration results. Split-beam and echo integration data, as well as water column depth, were processed in real-time (HTI; DEP, Version 3.5) and stored on a laptop for later data analysis.
**Data Processing**

Acoustic data were postprocessed using split-beam and echo integration analyses (Echoscape, version 2.10; HTI, Seattle, Washington). Split-beam analysis was used to determine the acoustic size (target strength, TS) of individual fish targets in decibels (dB). A sequence of echoes was accumulated according to a target tracking algorithm. Average TS and vertical position were stored in a database and referenced to elementary sampling distance units (ESDU) as defined by echo integration sampling resolution. For this study, data were divided into ESDUs with approximate horizontal \( \times \) vertical dimensions of \( 20 \times 2 \text{ m} \) for the first \( 0-2 \text{ m} \) of the water column, and \( 20 \times 1 \text{ m} \) for strata from \( 2 \text{ m} \) to the bottom. Using equations for clupeiform species, TSs obtained from the downward-oriented transducer (dorsal aspect) were converted to approximate fish length (Love 1977) and to wet weight using a generalized length–weight relationship for all forage fishes in the lake (J. Thompson, unpublished data). Echo integration provided measures of relative density that were converted to absolute volumetric density by scaling the integrated acoustic signal within an ESDU by the average target size within each ESDU. Volumetric densities were integrated throughout the water column to produce densities in two dimensions (expressed as fish/m\(^2\)). The database was incorporated into a geographical information systems (GIS) for data visualization and analyzed using S-PLUS (version 6.1; Insightful Corp.) for spatial structure and determination of abundance and biomass.

**Arithmetic and Geostatistical Analyses**

Two statistical procedures were used to calculate densities and estimate systemwide abundance and biomass. First, an arithmetic mean and variance of densities assuming identical and independent data (IID) were determined for the entire survey and then for each survey design (see below for description of survey designs). These summary statistics were extrapolated across the surface area of the lake and summed to produce systemwide abundance and variance of this estimate. The second procedure involved empirically modeling the spatial structure of the data using geostatistics. This technique involves three steps: spatial detrending, variogram analysis, and kriging.

**Spatial detrending.**—A regression-type model was used to explain large-scale spatial trends, such as those that may occur as a function of some larger spatial process (e.g., gradients in reservoir features resulting from heterogeneity in temperature or depth). For our analysis, a generalized additive model (GAM) was used with density as the response variable and spatial coordinates (latitude or “northing [km]” and longitude or “easting [km]”) as explanatory variables. This trend model served two functions. First, the GAM provides information on the large-scale spatial trend of density distributions throughout the reservoir. Second, the model removes the large-scale trend, or drift, in the mean density throughout the lake. Removing this drift is necessary to satisfy the assumptions of stationarity of the mean in spatial structure analysis (Cressie 1993). Output of this model includes (1) predictions of density, which are based on interpolations across the surface area of the lake and (2) residuals of the model fit, which are analyzed for small-scale spatial structure.

**Variogram analysis.**—The residuals from the GAM model are analyzed for small-scale spatial correlation using a variogram. The result is an empirical measure of correlation of pairs of points as a function of distance according to the equation

\[
\gamma_d = \sum \frac{(Z_{s+h} - Z_s)^2}{2N_h},
\]

where \( Z_s \) is the density at point \( s \), \( Z_{s+h} \) is the density at distance \( h \) from \( Z_s \), and \( N_h \) is the number of pairs of samples at distance \( h \) from each other. Based on the behavior of the empirical variogram, the variogram was modeled as a spherical theoretical variogram model using weighted, nonlinear least-squares fitting (Cressie 1993). The model contains three parameters (nugget, sill, and range) which have estimates and errors, and can be compared between and within models and data sets. The nugget corresponds to the variability that exists at a scale finer than that defined by the sampling resolution (<20 m in our case). The sill is the total or finite variance in the data set, similar to taking a global variance of a comparable, but uncorrelated, data set. Finally, the range is the distance beyond which points are independent. The range can be interpreted from a biological perspective as the maximum patch length observed in a survey.

**Kriging and abundance estimation.**—Finally, the variogram model is used as a best linear unbiased predictor to reconstruct the distribution of residual fish densities at unsampled points throughout the lake on a 20-m \( \times \) 20-m grid. Each grid cell contains a predicted residual density and an associated variance of the kriging prediction. To produce the
final density estimates, kriging predictions were added to the GAM trend predictions. These values were mapped in a GIS. Estimated total abundance was calculated by summing each of the kriging densities multiplied by kriging cell area. We determined the precision of the abundance estimates by approximating the variance of the abundance estimate through a summing of the kriging variances. Biomass was computed by multiplying the values for abundance by the average fish weight observed.

For the December 2001 data set, the geostatistical method was used in a four-way spatial distribution analysis to assess the performance of different sampling strategies as well as to assess spatial patterns in different regions of the reservoir. First, the entire data set was used to determine density and abundance. Second, a subset of the data set that represented only data collected along the longitudinal transects was analyzed. Third, only the cross-channel transects were analyzed. And finally, the reservoir was divided into three regions (Figure 1), and spatial analysis was performed on data collected within each region. The resulting abundance estimates for each region were then summed to provide a fourth estimate of systemwide abundance. Approximated variances and small-scale spatial structures characterized using the variograms were compared to assess the performance and differences in spatial patterns as resolved by each sampling and analytical method.

Resolving Vertical Structure

During the December 2001 survey, we observed two separate layers of targets in the water column, one layer at 2–3 m and a second layer at 11–16 m. We analyzed the spatial structure of these two layers by constructing separate variogram models for each. We compared the parameters of these variogram models in order to draw conclusions on how differences in patch structure between these layers could confound our approach for characterizing spatial patterns in two dimensions.

Results

July 2000 and December 2001 Lakewide Distribution and Abundance

Fish density measured in July 2000 was greatest in the upstream portion of the west arm of the reservoir (Figure 2). Vertical distribution was restricted to the top 10 m of the water column, a pattern that probably resulted from a significant thermo–oxycline that had developed in the system during the summer (J. Thompson, unpublished data). The GAM trend analysis explained 40% of the variability in the spatial distribution of density. The trend was only evident in the east–west direction, with more fish in the western region of the reservoir. Variogram analysis of the residuals from the GAM indicated a residual spatial structure. In general, densities exhibited a maximum patch length of 500 m (Figure 2, inset; Table 1). Prediction errors increased rapidly with distance from the sampled transect points, indicating no correlation between points greater than a 500-m distance (Figure 3). Densities ranged from 0 to 5 fish/m² across the reservoir. Estimated average densities using arithmetic and geostatistical procedures were 0.60 and 0.56 fish/m², respectively. The abundance estimate assuming independent data were $7.5 \times 10^6 \pm 4.61 \times 10^6$ (mean ± ap-
Table 1.—Comparison of abundance and density estimates assuming independent (IID) and spatially correlated (geostatistics) data for hydroacoustic estimates of forage fish in Badin Lake for July 2000 and December 2001 and various subsets of the December 2001 data. Variogram model parameters are also provided for figure insets.

<table>
<thead>
<tr>
<th></th>
<th>July</th>
<th>December</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>IID</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (fish/m²)</td>
<td>0.60</td>
<td>0.60</td>
<td>0.58</td>
<td>0.63</td>
<td>0.28</td>
<td>0.68</td>
<td>0.70</td>
<td>0.55</td>
</tr>
<tr>
<td>SDIID</td>
<td>0.61</td>
<td>0.69</td>
<td>0.72</td>
<td>0.61</td>
<td>0.46</td>
<td>0.77</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1.036</td>
<td>1.036</td>
<td>1.075</td>
<td>2.94</td>
<td>3.23</td>
<td>2.43</td>
<td>3.65</td>
<td>3.00</td>
</tr>
<tr>
<td>CV sampleIID×10⁶</td>
<td>101%</td>
<td>115%</td>
<td>122%</td>
<td>97%</td>
<td>161%</td>
<td>113%</td>
<td>93%</td>
<td></td>
</tr>
<tr>
<td>CV IID</td>
<td>3%</td>
<td>4%</td>
<td>4%</td>
<td>6%</td>
<td>11%</td>
<td>5%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Abundance (number of fish × 10⁶)</td>
<td>7.52</td>
<td>8.00</td>
<td>7.82</td>
<td>8.47</td>
<td>0.91</td>
<td>4.15</td>
<td>2.80</td>
<td>7.87</td>
</tr>
<tr>
<td>Variance (× 10⁶)</td>
<td>4.61</td>
<td>6.34</td>
<td>6.85</td>
<td>5.06</td>
<td>0.67</td>
<td>3.60</td>
<td>1.69</td>
<td>5.97</td>
</tr>
</tbody>
</table>

Geostatistics

<table>
<thead>
<tr>
<th></th>
<th>July</th>
<th>December</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Prediction mean (fish/m²)</td>
<td>0.56</td>
<td>0.57</td>
<td>0.56</td>
<td>0.72</td>
<td>0.32</td>
<td>0.57</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>Number of kriging points</td>
<td>31,342</td>
<td>33,448</td>
<td>33,448</td>
<td>33,448</td>
<td>8,075</td>
<td>15,325</td>
<td>10,048</td>
<td>33,448</td>
</tr>
<tr>
<td>SDGEO</td>
<td>0.28</td>
<td>0.49</td>
<td>0.47</td>
<td>0.46</td>
<td>0.28</td>
<td>0.45</td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td>CVGEO</td>
<td>10%</td>
<td>29%</td>
<td>29%</td>
<td>15%</td>
<td>17%</td>
<td>28%</td>
<td>29%</td>
<td></td>
</tr>
<tr>
<td>Abundance (number of fish × 10⁶)</td>
<td>7.05</td>
<td>7.07</td>
<td>7.30</td>
<td>9.69</td>
<td>1.03</td>
<td>3.53</td>
<td>2.79</td>
<td>7.35</td>
</tr>
<tr>
<td>Approximate variance (× 10⁶)</td>
<td>0.83</td>
<td>2.10</td>
<td>2.10</td>
<td>2.60</td>
<td>0.22</td>
<td>0.90</td>
<td>0.87</td>
<td>1.98</td>
</tr>
</tbody>
</table>

Variogram parameters

| Range   | 0.51          | 0.78         | 1.07  | 0.55  | 0.47  | 0.46  | 0.18  |       |
| Still   | 0.08          | 0.24         | 0.22  | 0.21  | 0.08  | 0.20  | 0.25  |       |
| Nugget  | 0.03          | 0.07         | 0.06  | 0.10  | 0.02  | 0.04  | 0.05  |       |

*a Coefficient of variation (100 × SD/mean).

Figure 3.—Prediction standard errors (S.E.) for kriged density distribution of forage fish in July 2000. Values range from 0.19 to 0.30 according to the gray scale and increase rapidly with distance from sampled transects.

proximated variance) fish. The abundance estimate using the kriging approach was similar to the arithmetic estimate, but the variance was fivefold lower (7.05 × 10⁶ ± 8.30 × 10⁵; Table 1). The sample standard deviation for the geostatistical method was also less than half that calculated assuming IID. Average observed target size from the downward-oriented data for July 2000 was −50.9 dB. Using Love’s (1977) equation for our frequency of 200 kHz,

\[
TL = 10^{(TS - 64.1)/20},
\]

where TL is the estimated total length and TS is the average observed TS, we estimated an average size of 45 mm TL. We converted this length to an average wet weight (WW) using a generalized equation for all forage species in the reservoir (J. Thompson, unpublished data), namely,

\[
WW = 7 \times 10^{-6}(TL)^{3.11}.
\]

This resulted in an average wet weight of 0.7 g. By multiplying this average weight by the abundance estimate from the arithmetic average and
Fish were concentrated in the western arm of the reservoir in December 2001, with higher densities in the middle third of the arm (Figure 4). Two horizontal bands of targets were apparent in the reservoir. Targets were observed in the 2–3-m and the 11–16-m depth strata. The water column was vertically integrated to produce two-dimensional densities for the analysis; however, below we comment further on the spatial structure within these two strata. The GAM analysis explained 8% of the variability in density distribution. There was a significant positive trend toward the western and northern regions of the reservoir. Maximum patch size was 780 m. The most distinct patches were in the western arm as observed in July; however, there were also distinct patches observed in other regions of the reservoir (Figure 4). A complete description of regional spatial patterns is included below. As with July, prediction errors for the kriged density distribution increased rapidly with distance from the sampled transects. Average densities calculated assuming IID and spatially correlated data were 0.60 and 0.57 fish/m², respectively. Abundance estimates in December were $8.00 \times 10^6 \pm 6.34 \times 10^6$ and $7.07 \times 10^6 \pm 2.10 \times 10^6$ for the arithmetic and geostatistical methods, respectively (Table 1). The approximated variance estimate using the kriging approach was one-third of the variance estimate assuming independent data (Table 1). Average target sizes observed in December were 48.9 dB, or 57 mm TL (1.3 g wet weight). The systemwide biomass estimation was 11,200 and 9,901 kg of forage fish for December 2001 using the arithmetic average and geostatistical abundance estimates, respectively.

**December 2001 Sampling and Regional Comparisons**

A comparison of sampling and analytical methods used on the December 2001 data set resulted in differences in the characterization of spatial patterns in the reservoir as well as in the magnitude and precision of the abundance estimates using the geostatistical method. Using only the longitudinal transects produced a reasonable abundance estimate when compared with the results generated from the analysis of the entire data set. The GAM analysis explained 10% of the variability in distribution. The spatial structuring was different than what was observed in the entire data set, with the maximum patch length extending to just over 1 km compared with 780 m (Table 1; Figure 5, inset). Reconstruction of the density distribution for these longitudinal transects produced similar results in spatial distribution and abundance estimates when compared with the patterns generated using the entire data set, with fish accumulated in the upper reaches of the western arm of the reservoir (Figure 5). Approximated variance estimates were equivalent using the longitudinal transects when compared with results generated using the entire data set (Table 1).

Density distributions were markedly different when reconstructed using only the cross-channel transects (Figure 6). A small portion of the variability was explained in the GAM trend analysis (8%), which predicted high densities of fish in the
western arm. The nugget parameter of the variogram was also the highest of all analytical approaches, indicating less variability was explained by the spatial analysis. Average densities were highest using the kriging method with an average of 0.72 fish/m², compared with 0.63 fish/m² assuming IID. The cross-channel sampling approach resulted in the poorest precision with respect to approximated variances of the abundance estimates (Table 1). Abundance estimates were high when compared with the other three methods. The systemwide abundance estimates generated using only the cross-channel transects were $8.47 \times 10^6 \pm 5.06 \times 10^6$ and $9.69 \times 10^6 \pm 2.60 \times 10^6$ fish using the arithmetic and geostatistical methods, respectively (Table 1).

Analysis of the data set by region resulted in distinct spatial patterns in each arm of the reservoir (Figure 7). Large-scale trends in each of the regions were quite different as discerned in the GAM analysis. The large-scale trend explained 12, 22, and 0.1% of the spatial variability in density in regions 1–3, respectively. Region 3 had the shortest maximum patch length at just less than 200 m, followed by region 2 and region 1 (456 m and 470 m, respectively). Assuming either IID or correlated data, the highest densities were observed in region 3; however, owing to a greater surface area, highest abundance was observed in region 2, which held just under 50% of the forage fish in the reservoir (Table 1; Figure 7). Region 1 had the lowest density and, owing to the small surface area, held the fewest fish. Integration of the regional abundance estimates produced systemwide estimates using
FIGURE 7.—Reconstruction of density distributions for forage fish sampled in December 2001, analyzed for each region separately. Variogram models for each region are given; model parameters are available in Table 1.

FIGURE 8.—Example of vertical distribution of forage fish for July 2000 (top) and December 2001 (bottom) as observed along a similar region of Badin Lake. Forage fish density is proportional to the gray-scale shading.

Observations on Vertical Distribution Patterns

Fish targets were observed in two distinct strata in the December sampling. The upper stratum was observed between a depth of 2 and 3 m, and the lower stratum was between a depth of 11 and 16 m (Figure 8). Target sizes of these two strata were similar (J. C. Taylor, unpublished data); however, there were differences in the spatial structure in these two layers. The upper stratum had a shorter maximum patch length (Figure 9). The sill parameter in the variogram model for the upper stratum was also two orders of magnitude higher than that for the lower stratum (Figure 9).

Discussion

The density distribution of forage fish in Badin Lake exhibited significant large- and small-scale spatial variation regardless of the month sampled or the sampling strategy used. In general, large-scale trends were evident in July and December, with more fish observed in the upstream region of the western arm (region 2) of the reservoir. Higher densities of age-0 gizzard shad in upstream reaches of reservoirs have been attributed to spawning (Tisa and Ney 1991), but seasonal or system-specific variability in these patterns have also been observed (Michaletz and Gale 1999). We did observe higher densities in the upstream regions of the eastern arm (region 3) as well, although the pattern was far less pronounced when compared to the trend in the east–west direction. Habitat features, such as temperature and prey in these upstream regions, may have been favorable throughout the year as we observed high densities in December, which is outside the period of reported

Arithmetic and geostatistical methods of $7.87 \times 10^6 \pm 5.97 \times 10^6$ and $7.35 \times 10^6 \pm 1.98 \times 10^6$, respectively (Table 1). These estimates were comparable to the analysis of the entire data set in magnitude, with slightly improved precision (Table 1).
spawning by shad (Irwin and Bettoli 1995; Hale 1996).

The lower variability in density distribution explained by the GAM trend analysis, particularly in the cross-channel transect design (8%) and region 3 (0.1%), coupled with the shorter maximum patch sizes, can be attributed to the presence of small, dense patches of fishes observed along transects. Increased heterogeneity in the density distribution can decrease the strength of the observed trend or gradient in forage fish distribution. For both months, densities along each transect exhibited autocorrelation even after the removal of the large-scale trend using a GAM model. This autocorrelation in the residuals of the GAM models violates the requirement of independence in the data points, which renders standard arithmetic means and estimates of variance inappropriate and calls for an explicit analysis of the small-scale spatial structure using geostatistics (Cliff and Ord 1981; Cressie 1993; Rivoirard et al. 2000).

Spatially autocorrelated data are not unexpected for clupeoid species during the daytime based on the results of a number of field studies using hydroacoustics (Blaxter and Hunter 1982; Pitcher 1986; Freon and Misund 1999). Even during the nighttime sampling in the present study, there was residual spatial structure in the lake. Other studies on forage fish populations in reservoirs and other clupeoids in coastal systems have also found small-scale spatial patterns at night (Schael et al. 1995; Taylor and Rand 2003). Using a patch detection algorithm, Schael et al. (1995) found patches that were 50 m to well over 500 m in length, depending on the level of data filter used (an index of the density within patches). Our variogram analysis, while not explicitly identifying patch numbers or individual patch sizes, provides a measure of maximum patch sizes that existed during our sampling. The ranges for our variogram models varied from 0.2 to 1.1 km. Shorter ranges were observed lakewide in July 2000 compared with ranges observed lakewide during December 2001; however, we did find short ranges in region 3 in December. Schael et al. (1995) also observed small, but very dense patches in certain regions of Lake Norman (another reservoir in North Carolina) in the summer when compared with October samples. Abundance estimates were less precise in December 2001 than in July 2000, likely due to increased heterogeneity in the distribution pattern that was encountered in regions of the reservoir (e.g., region 3) in December.

A comparison of the spatial correlation in the entire data set with that of the longitudinal and cross-channel subsets resulted in differences in spatial patterns and in the precision of abundance estimates. The analysis including only the longitudinal transect data appeared to capture much of the large- and small-scale spatial variability when compared with the results from the analysis of the entire data set. In contrast, the cross-channel transects resulted in the highest abundance and variance estimates of the sampling designs analyzed using the geostatistical approach. Using the entire data set, high densities were observed in the region of the most westward cross-channel transect; however, our longitudinal transect sampling indicated densities decreased further up in the western arm. This decline was not resolved by the cross-channel transect sampling, and thus explains the resulting overestimate using the cross-channel transect sampling approach. Another indication of increased uncertainty in the estimates from the cross-channel transect design is a comparison of the nugget parameter (an index of measurement error or spatial
uncertainty occurring at scales smaller than that defined by the sampling resolution of our hydroacoustic gear) and the sill (the total variance in the data set). In the case of the cross-channel transect, the nugget was over 47% of the sill, indicating only a small portion of additional uncertainty was explained using the variogram analysis (Isaaks and Srivastava 1986). We recognize that the effort of the cross-sectional transects was much lower when compared with that of the longitudinal transects. However, sampling in the very upstream region of the reservoir would not have been feasible using parallel, cross-channel transects. In the interest of improving the precision of systemwide estimates of abundance and biomass in reservoir systems while minimizing effort, we would suggest focusing effort on longitudinal transects based on the results we report here for Badin Lake. A comparison of effort in our December 2001 sampling strategies indicated a 40% reduction in operator-hours by using only the longitudinal transect design when compared with sampling both the longitudinal and cross-channel transects.

Despite the relatively poor results in the kriging predictions of the cross-channel transects, the longer range in the longitudinal data (1.07 km) and the short range in the cross-channel data set (0.54 km) could indicate anisotropy in the data; that is, patches of fish may be characterized better as an oval or ellipse whose long axis may run parallel to the longitudinal axis of the reservoir. While the comparison of the two transect sampling designs in this study suggests sampling could be optimized by relying solely on data collected along longitudinal transects, care should be taken in the biological interpretation of the range parameter based solely on longitudinal sampling. Instead, further analysis of the anisotropy is necessary to gain a better understanding of density distribution across the entire reservoir, especially with respect to prey patch characteristics in two or three spatial dimensions. The orientation of the different regions of the lake complicates an analysis of anisotropy, and it is likely that this anisotropy should be analyzed in each region or smaller subregions (Simard et al. 1992; Rivoirard et al. 2000).

The regional analysis provided an improvement in the precision of the systemwide abundance estimate when compared with the results from the analysis of the entire data set. In addition, using the regional analysis provided even more information on the spatial patterns of distribution throughout the lake, such as region-specific scale-of-prey distributions that may influence trophic dynamics and predator habitat selection. Spatial structure in the eastern arm was dominated by small-scale patterns, possibly owing to more heterogeneous conditions resulting from the multitude of smaller tributaries within this arm (Van Den Avyle et al. 1982; Siler et al. 1986).

This study, along with previous analyses of hydroacoustic surveys for prey fishes in reservoirs, uses data that are typically integrated through the water column in the vertical dimension (Degan and Wilson 1995; Cyterski et al. 2003). In July, vertical distribution was restricted to the top 10 m of the water column due to a significant thermocline. This vertical concentration of fishes in monomictic reservoirs is common under summer seasons when stratification can reduce suitable habitat volume primarily through a reduction of available oxygen for metabolism (Coutant 1985; Schael et al. 1995). The water column was well mixed during the December 2001 sampling period. Distinct vertical separation of forage fishes was evident during this sampling period, with targets observed in the top 3 m of the water column and another group occupying habitat 11–16 m below the surface (Figure 8). The range of autocorrelation in the top layer was shorter than in the deeper strata. In addition, the finite variance in the top layer was 44 times that in the deeper strata. Loosely interpreted, this suggests patches were more dense and smaller in the top stratum when compared with the bottom stratum. Therefore, finer scale spatial patterns that were evident in each stratum in December may have been masked by the vertical integration of the three-dimensional data necessary to generate our estimates of biomass expressed in two dimensions (e.g., fish or g/m²). As a result, we caution the use of this depth-integrated approach for density and abundance estimation in studies focused on ecological interactions that may be occurring at a specific depth stratum. We were unable to collect samples of fishes in water deeper than 10 m given the drop length of the purse seine used in this study. While we found target strengths to be similar between the two strata, it remains unclear what species are residing in this deep layer. Vertical structure was also observed during certain seasons in Lake Norman, a nearby reservoir in North Carolina (Schael et al. 1995). Further analysis of these data in three-dimensions may provide insight into the seasonal patterns of distribution of prey species in reservoir systems and help address questions of habitat selection exhibited by piscivorous fish.
Acknowledgments

We thank B. Degan, J. Rice, and S. Waters for assistance in the field, and N. Bez, D. Degan, E. Gardner, S. Johnston, P. Nealson, P. Rivoirard, and P. Sullivan for technical advice. The manuscript was significantly improved through discussions with L. Avens, D. Coughlan, D. Degan, and A. M. Mueller, and from helpful comments provided by two anonymous reviewers. This project was supported in part by a North Carolina Wildlife Resource Commission grant (Cooperative Projects with Universities Number P 68-O2 to P.S.R., J. Rice, and J. Hightower), an EPA Science to Achieve Results Fellowship to J.C.T., and a National Science Foundation predoctoral fellowship to J.S.T.

References


